Observing System Development and UQ in a Parallel Bayesian Framework: Applications for Weather, Clouds, Convection, and Precipitation

27 May 2021

Presenter and PI: Derek J. Posselt1

Team Members: Brian Wilson[1], Rachel Storer[2], Derek Tropf[1], Vishal Lall[1], George Duffy[1], Matt Lebsock[1], Simone Tanelli[1]

Program: AIST-18

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA

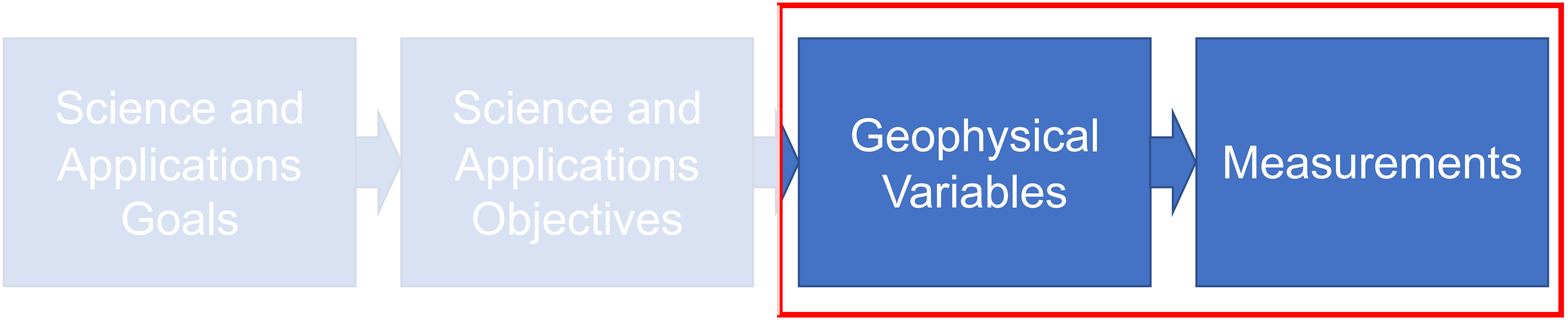
[2]University of California, Los Angeles, CA

One of these talks is not like the others…

- Technologies in this session provide information highly relevant for weather
- This talk describes *information technology* for weather formulation
- Rapidly expanding trade space – what is needed?
  Which instruments? Combinations/constellations? Accuracy?
- We have developed a system that is designed to more thoroughly and efficiently explore the science trade-space for new missions.
- It is flexible, parallelizes over diverse architectures, and includes several robust techniques with which to measure uncertainty.
- When combined with tools (e.g., TAT-C) that assess sampling needs (orbits, swaths, etc) it is now possible to evaluate a much larger number of options.

Scientific Challenge

- Clouds and precipitation are central to weather and climate
- After decades of space-borne measurements, *key processes are still missing*
- Goal: design a new observing system (e.g. ACCP*)
  - Address specific science objectives
  - Consider the vast array of possible measurements
  - Rigorously quantify uncertainties



*Aerosol, Clouds, Convection, and Precipitation https://science.nasa.gov/earth-science/decadal-accp

Technical Challenge

- The design trade-space is *large* and clouds are *diverse*
- The dimensionality of the design problem is *immense*
  - Multiple different geophysical scenarios (different cloud types)
  - Diversity of measurement types (active, passive, single-point, distributed)
  - Multiple sources of uncertainty (instrument noise, forward models, ambiguity)

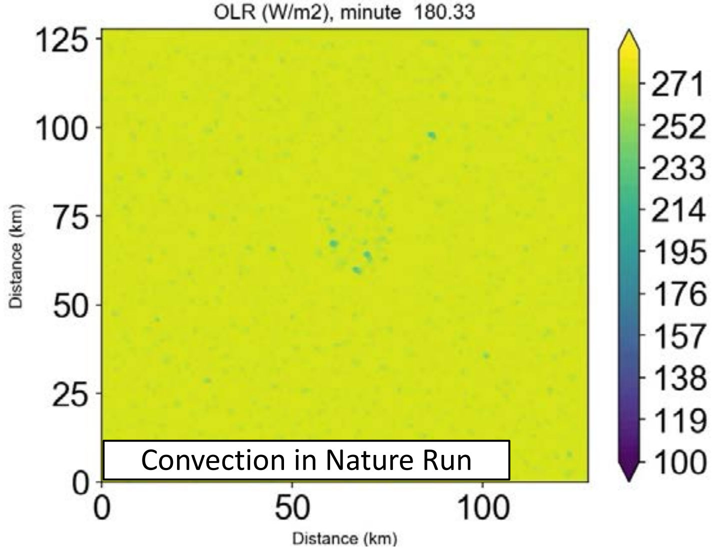
For each geophysical variable → For each cloud type → For each combination of measurements → Considering all sources of uncertainty → Determine whether measurements meet mission requirements

Solution: Accelerate OSSEs

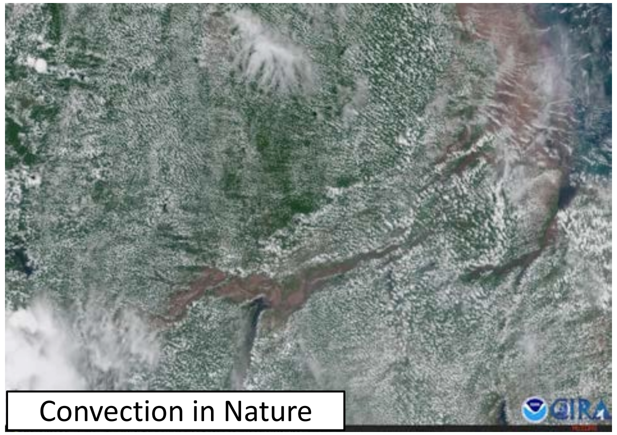
Any **observing system simulation experiment** (OSSE) requires at least four components:

1. Nature run: Realistically represent the real world

Simulation of Convection over Amazon



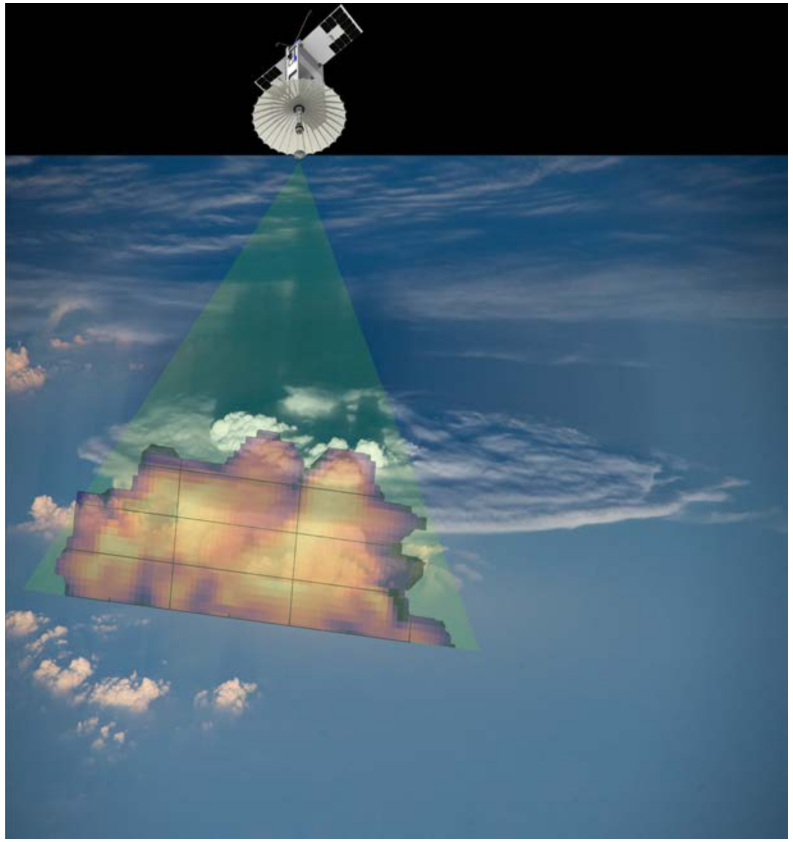
GOES-16 Observations of Convection over Amazon



Solution: Accelerate OSSEs

Any **observing system simulation experiment** (OSSE) requires at least four components:

1. Nature run: Realistically represent the real world
2. Instrument simulators: Synthetic measurements

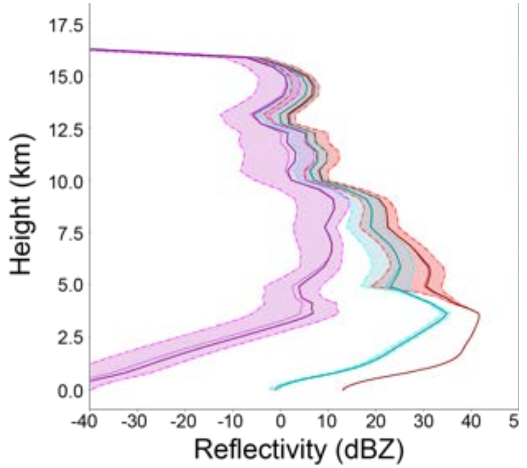


Solution: Accelerate OSSEs

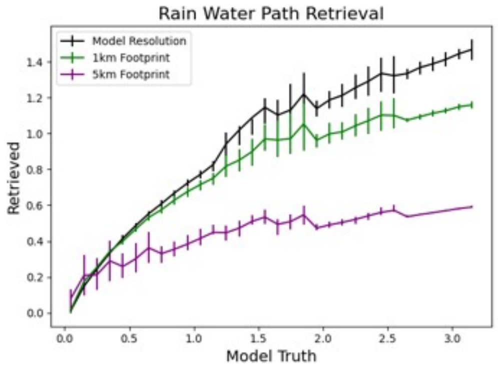
Any **observing system simulation experiment** (OSSE) requires at least four components:

1. Nature run: Realistically represent the real world
2. Instrument simulators: Synthetic measurements
3. Quantify uncertainty: Sources of noise and error

Measurement Uncertainty



Geophysical Variable Uncertainty

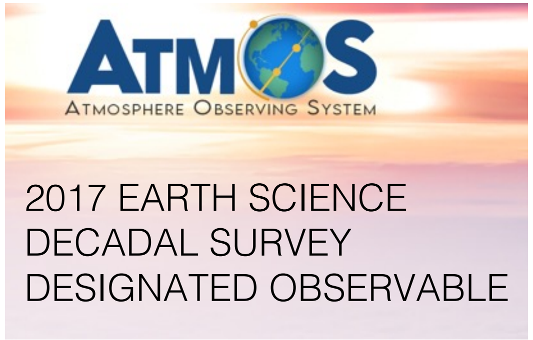


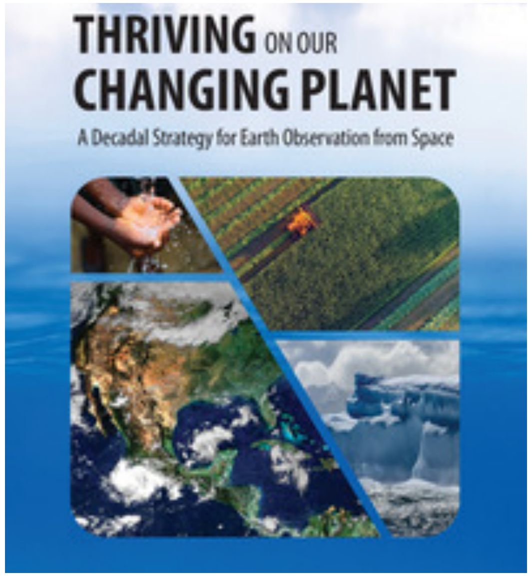
Solution: Accelerate OSSEs

Any **observing system simulation experiment** (OSSE) requires at least four components:

1. Nature run: Realistically represent the real world
2. Instrument simulators: Synthetic measurements
3. Quantify uncertainty: Sources of noise and error
4. Assess impact*: Did observations meet science and applications goals and objectives?

*NWP (weather forecast OSSE) is just one example of impact. OSSEs must grow to encompass advances in knowledge and traceability to applications.





Parallel OSSE Toolkit for Mission Design

OSSE Components

Nature Runs

Large Eddy Simulations
Cloud Resolving Models
Global Simulations

Instrument Simulation

Radar
Passive Microwave
(Extensible via pluggable containers)

Bayesian Retrievals

Optimal Estimation
Ensemble Kalman Filter
Markov chain Monte Carlo

ParMAP: Flexible Parallelism

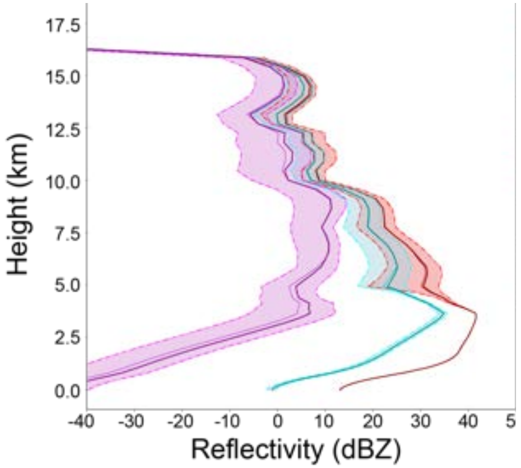
Standalone Workstation

Clusters and HPC

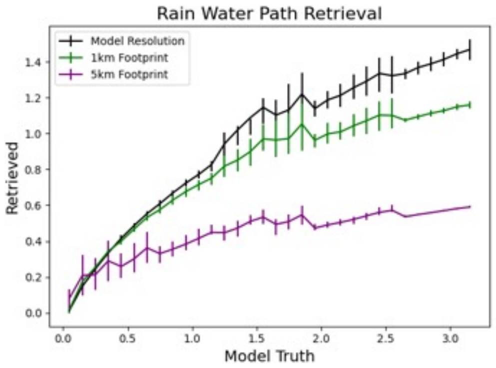
Cloud Computing

Uncertainty Analysis

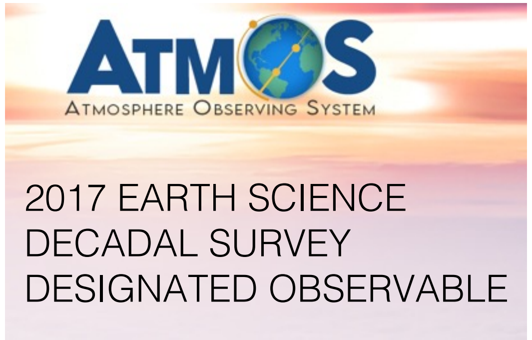
Measurement Uncertainty



Geophysical Variable Uncertainty

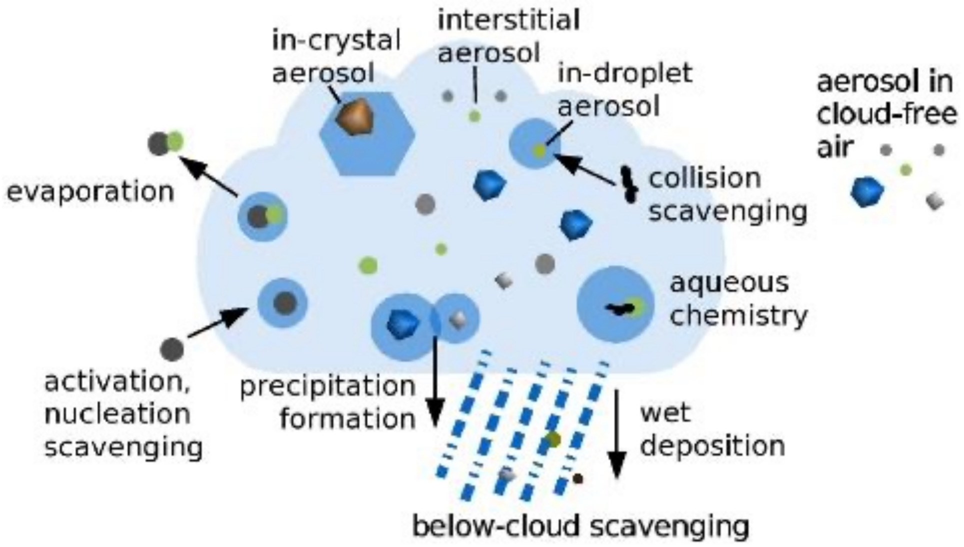


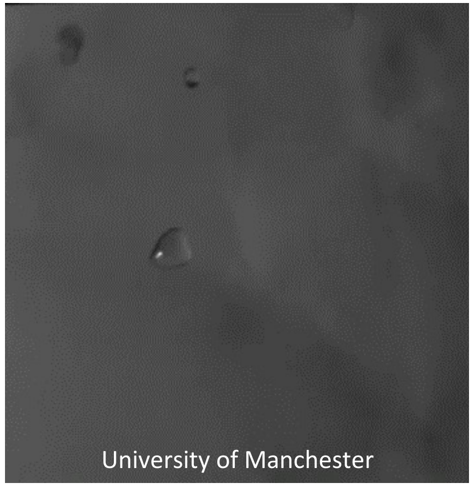
Mission Design Decisions



Example 1: Uncertainty Inherent in Clouds

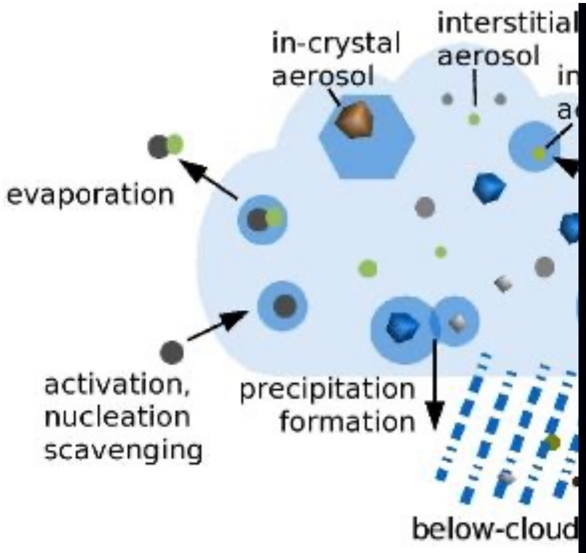
- Accurate estimates of cloud properties and evolution are important
  - Precipitation
  - Atmospheric dynamics
  - Earth's radiative balance
  - Chemical reactions
- Many processes of interest are governed by cloud microphysics:
  - Phase change, collisions, etc
  - Particle size, number, and shape



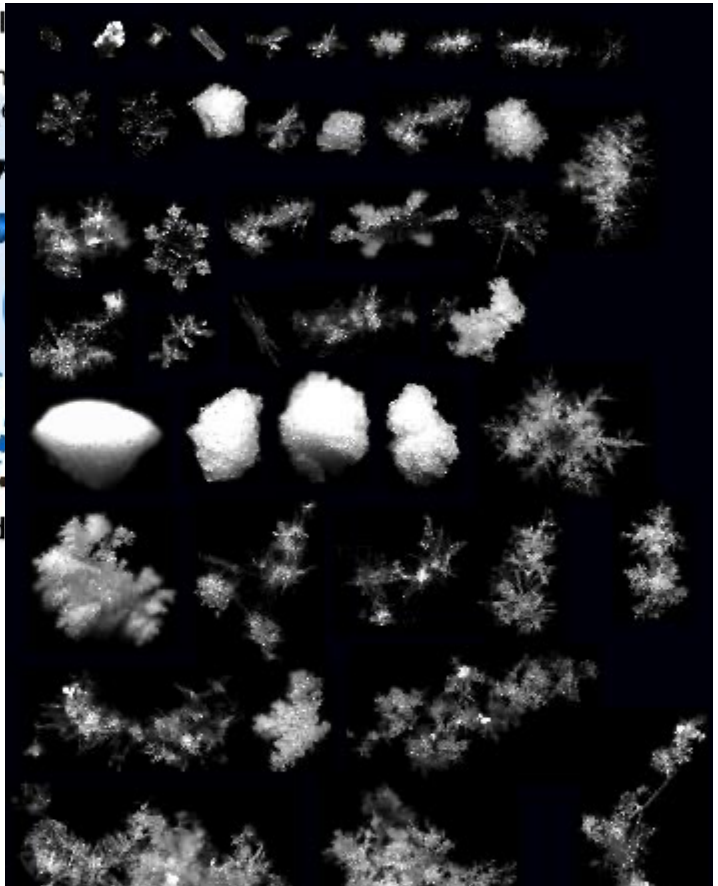


Example 1: Uncertainty Inherent in Clouds

- There is a diversity of available remote sensing measurements
- All are sensitive to some degree to cloud microphysics
- What are the measurement requirements for successfully observing cloud properties and processes?



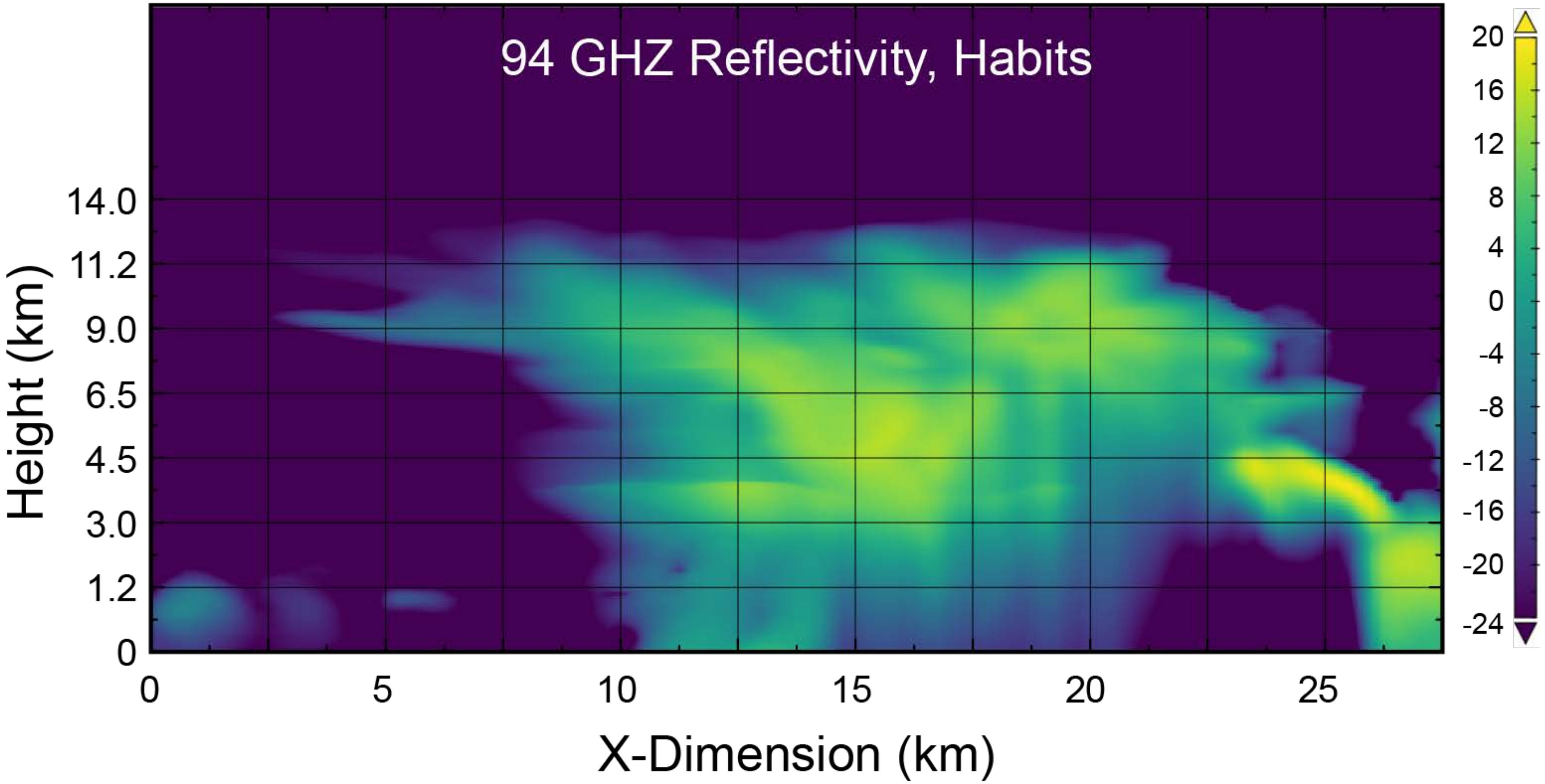
University of Manchester



Tim Garrett, Univ. Utah
http://www.inscc.utah.edu/~tgarrett/Snowflakes/Gallery/

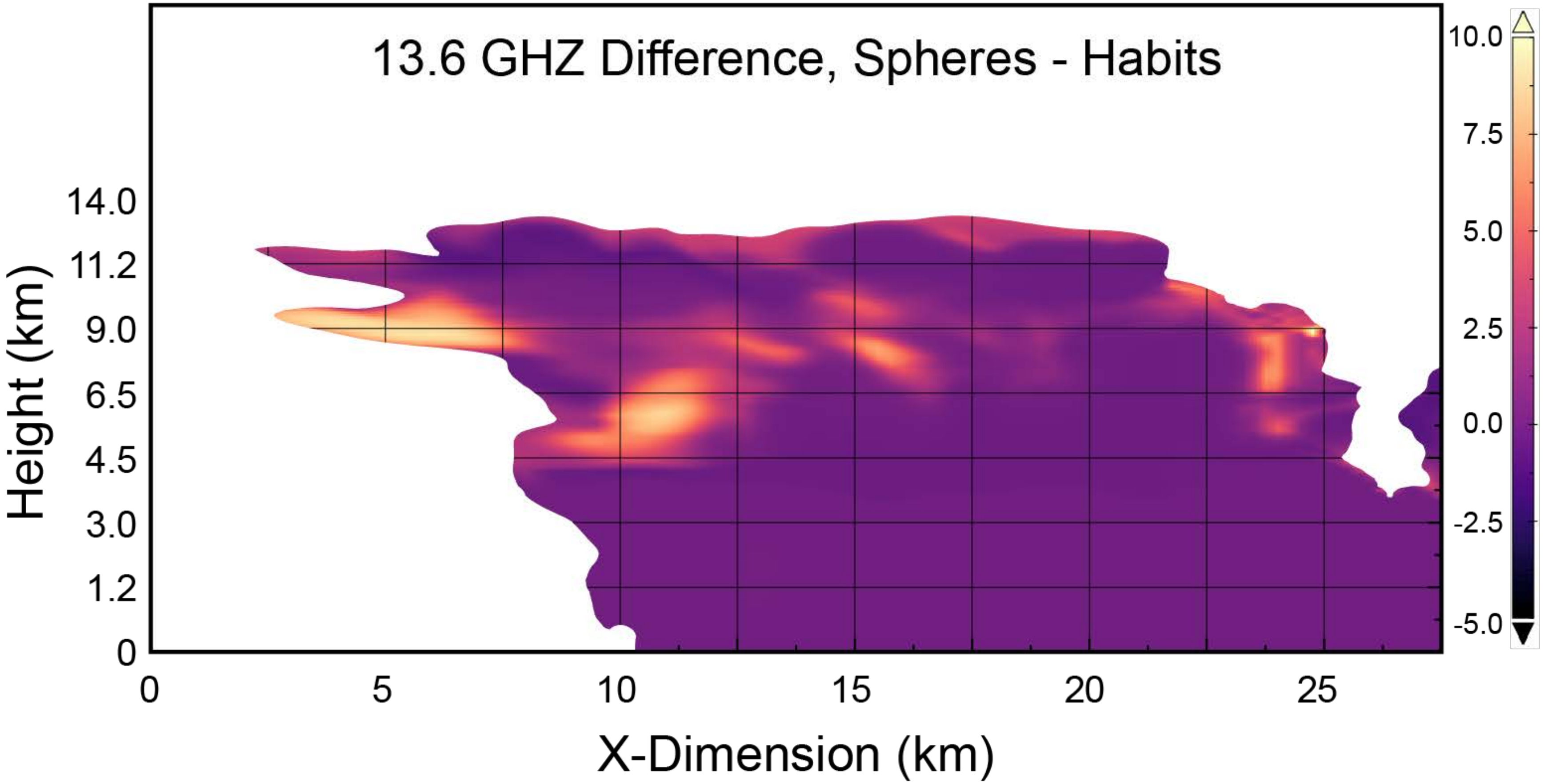
Example 1: Uncertainty Inherent in Clouds

- Quick example: run a radar simulator using two different sets of ice shapes (spheres vs multiple habits)



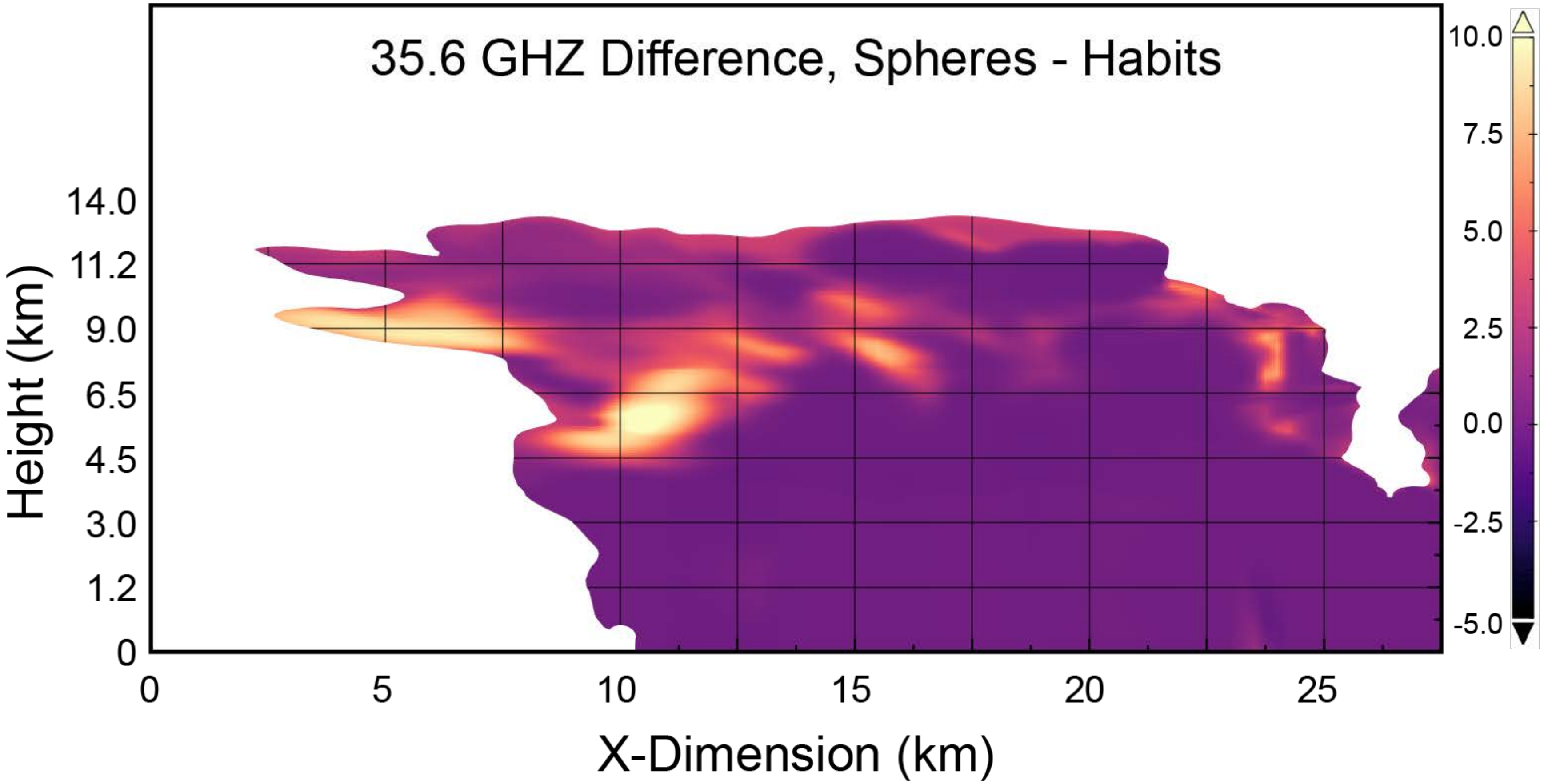
Example 1: Uncertainty Inherent in Clouds

- Quick example: run a radar simulator using two different sets of ice shapes (spheres vs multiple habits)



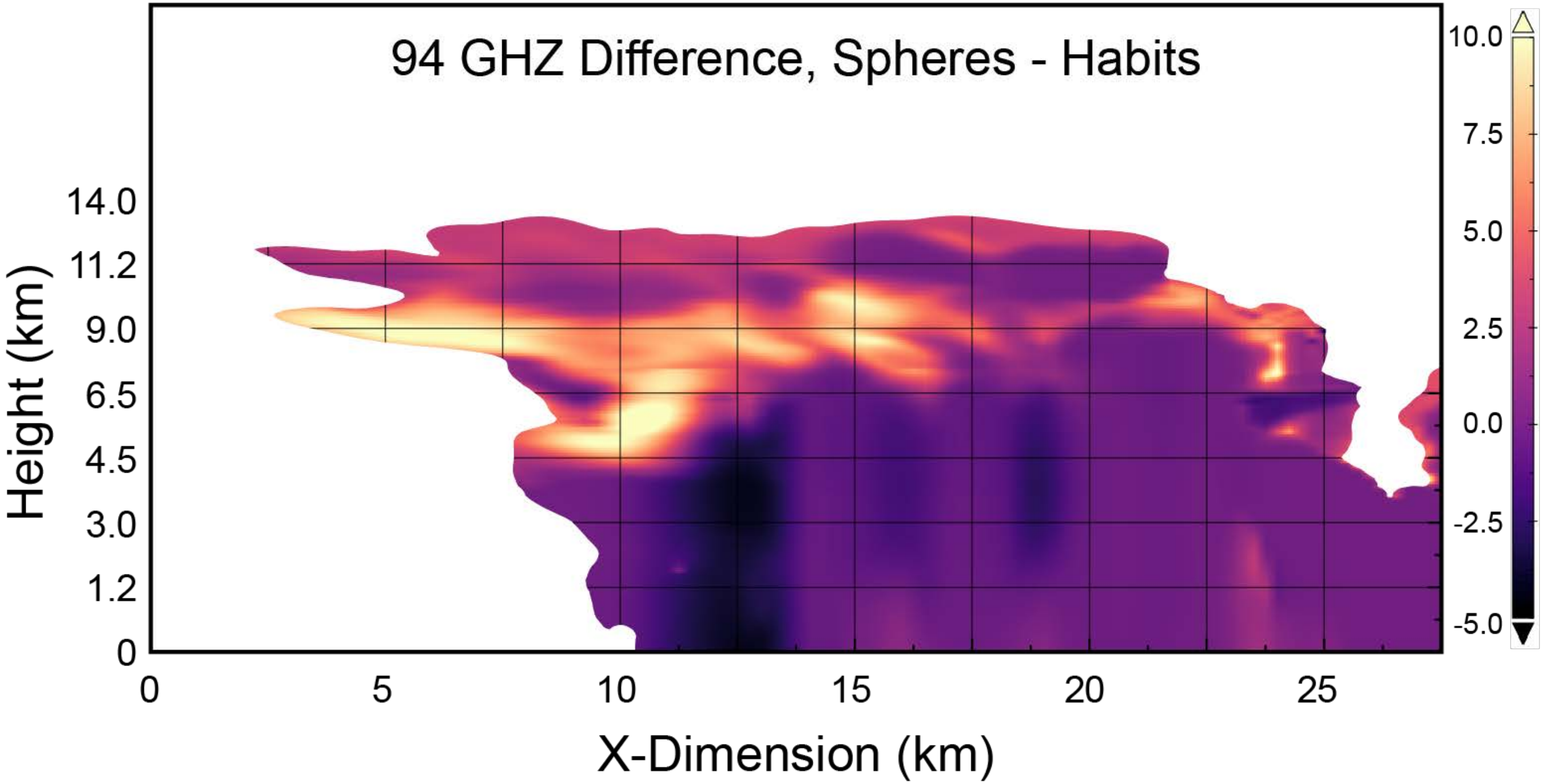
Example 1: Uncertainty Inherent in Clouds

- Quick example: run a radar simulator using two different sets of ice shapes (spheres vs multiple habits)



Example 1: Uncertainty Inherent in Clouds

- Quick example: run a radar simulator using two different sets of ice shapes (spheres vs multiple habits)



Example 1: Uncertainty Inherent in Clouds

Experiment Configuration:

- 2 input model profiles
- 3 radar frequencies (Ku, Ka, W)
- 5 uncertain parameters, 11 possible values each
- 2 x 3 x 11^5 = 966,306 forward model runs

Inputs:

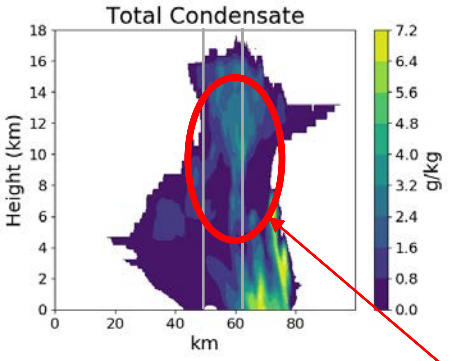
- Nature run profiles
- Range of uncertainty

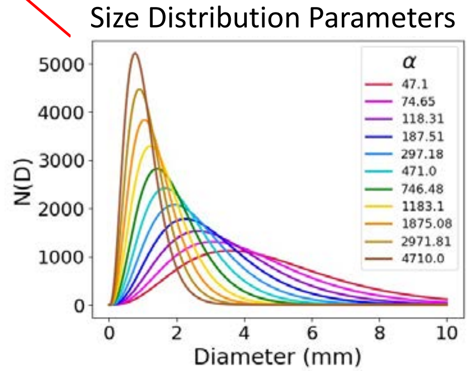
Outputs:

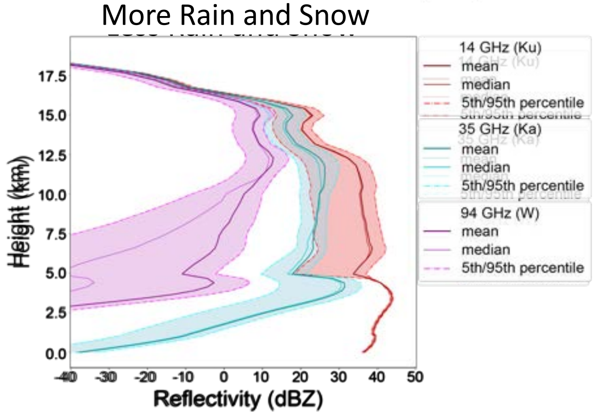
- Ensemble of possible radar profiles for each input model profile and frequency
- Improved understanding of uncertainty in radar observations of convection

~10^6 Forward Model Runs

60+ hours sequential
2 hours parallel
(40 cores) 36x speedup



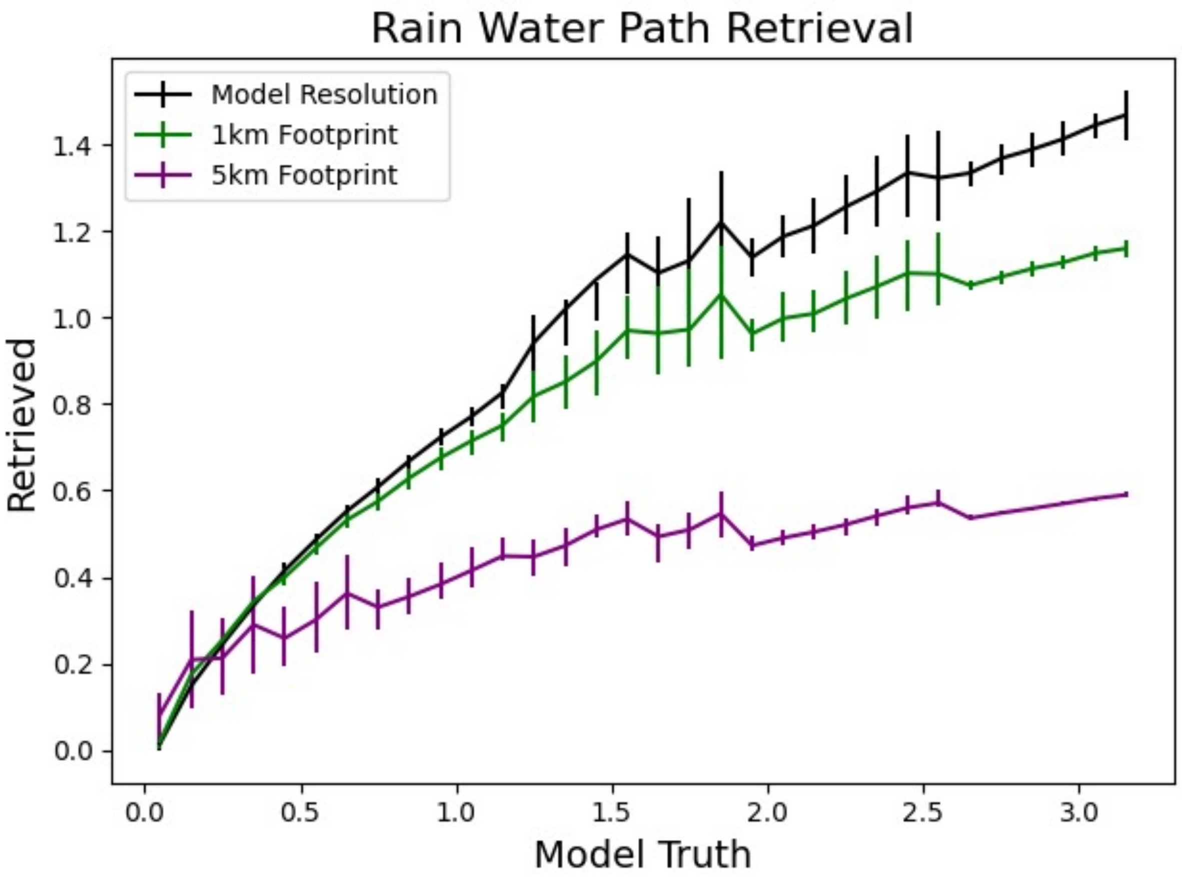




Posselt et al. 2021 (IGARSS)

Example 2: Shallow Cloud Retrieval

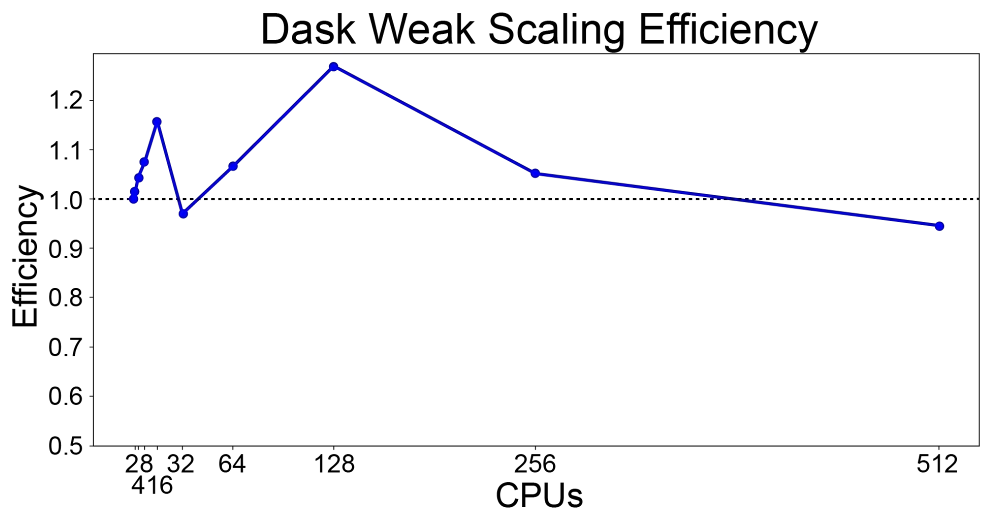
- Shallow convection is crucial for climate (hydrologic cycle and cloud-radiation feedbacks)
- Rain retrievals are challenging: sensitive to radar design parameters (sensitivity, footprint, surface clutter)
- Constructed an optimal estimation (Bayesian) retrieval based on the CloudSat algorithm
- Conducted an initial test of retrieval uncertainty using 6000 shallow rain profiles from nature run

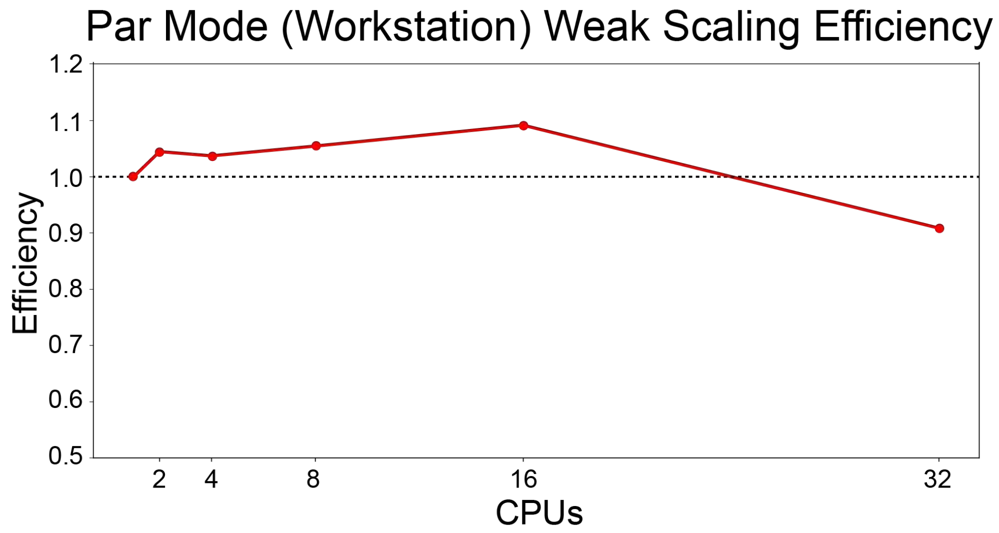


Parallel OSSE System (ParOSSE) Performance

- Sensitivity and retrieval experiments are embarrasingly parallel (can be done nearly independently)
- ParMAP library makes ParOSSE deployable on a single machine (Par), cluster (Dask), and AWS Lambdas
- Our initial tests have indicated excellent scaling efficiency*

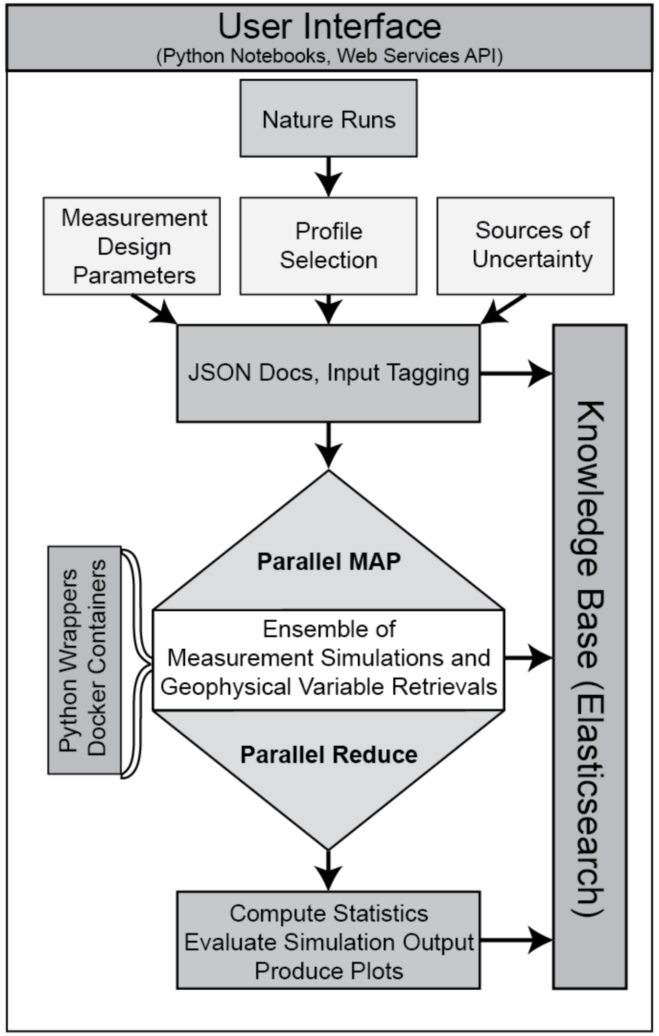
*Efficiency > 1 is due to I/O limitations with a single CPU





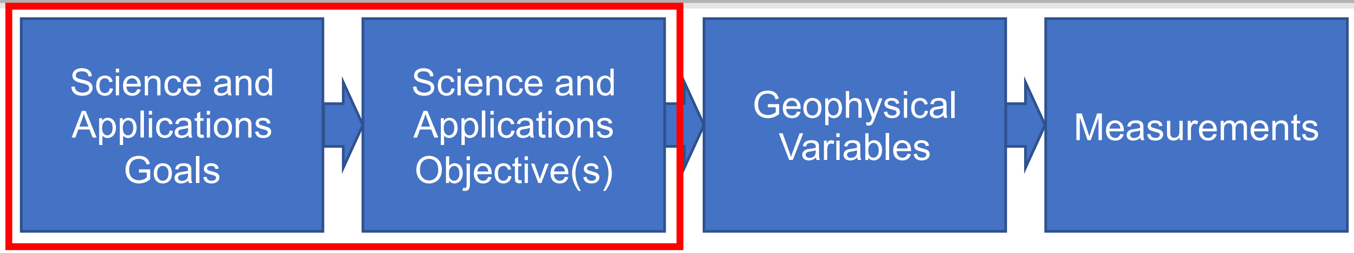
ParOSSE Capability to Date

- Pluggable nature runs and instrument simulators enable a wide range of trade space studies
- Flexible parallelism enables experiments on diverse architectures and more thorough exploration of uncertainty in measurements and retrievals
- Have implemented various sensitivity analysis techniques
  - Method of Morris, Sobol sensitivity, Monte Carlo, grid search
- Retrievals can utilize several Bayesian methodologies
  - Optimal estimation, MCMC, ensemble Kalman filter, Gamma-Inverse Gamma filter



Future Directions: Span the SATM?

Can we quantify ability to meet science and applications goals and objectives?



Science:

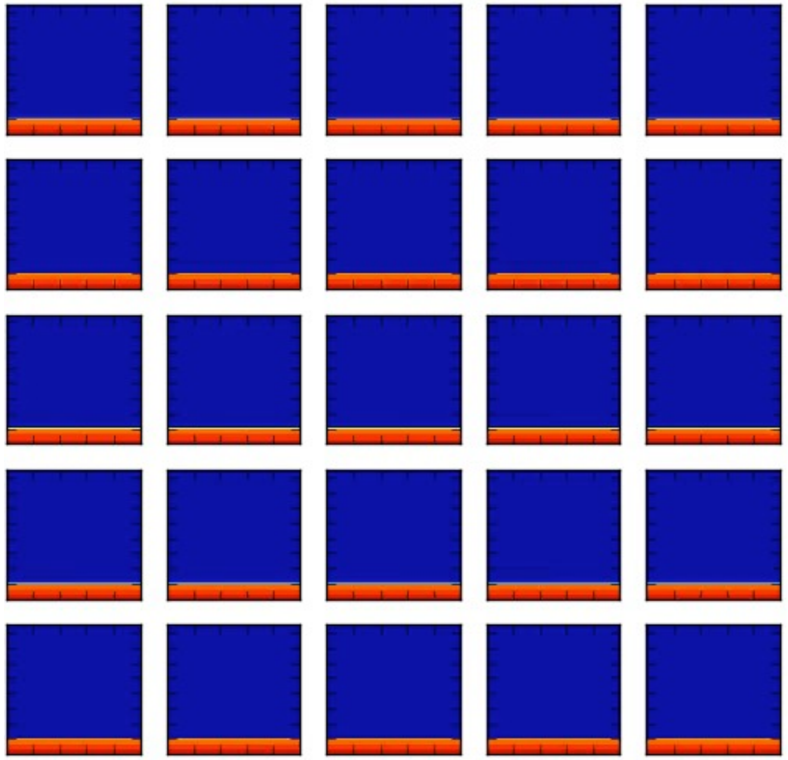
- Quantify state of knowledge – sources of uncertainty and relevant variables?
- Models as a laboratory, and ensembles as the tool.
- ParOSSE is flexible - spawn ensembles of *process simulations* and assess reduction in uncertainty (metrics from information theory, ensemble forecasting, etc)

Applications:

- Map from GV uncertainty to uncertainty in stakeholder quantities of interest (e.g., rainfall duration and intensity vs. needs of reservoir managers)
- Note: NOAA's ASPEN system considers a large database of user-defined requirements and then quantifies observing system effectiveness by inputting expected GV uncertainty.

Example: Convection-Environment Interaction

- Which observations are necessary to improve state of knowledge of convective storms?
- First: determine which are the most important control variables
- How? Models as a laboratory
- This is a small number of runs of one case, each with a slightly different environment
- Can we scale up to many types of convection in many different environments?
- ParOSSE's flexible configuration makes this straightforward



Cross-section through ensemble of 25 simulations of deep convection, showing transport of pollution from the boundary layer upward into the free troposphere.

References

- Maahn, M., D. D. Turner, U. Lohnert, D. J. Posselt, K. Ebell, G. G. Mace, and J. M. Comstock, 2020: Optimal Estimation Retrievals and Their Uncertainties: What Every Atmospheric Scientist Should Know. Bull. Amer. Meteor. Soc., 101, E1512–E1523, https://doi.org/10.1175/BAMS-D-19-0027.1
- Posselt, D. J., T. S. L'Ecuyer, and G. L. Stephens, 2008: Exploring the Error Characteristics of Thin Ice Cloud Property Retrievals Using a Markov Chain Monte Carlo Algorithm. J. Geophys. Res., 113, D24206, doi:10.1029/2008JD010832.
- Posselt, D. J., and T. Vukicevic, 2010: Robust Characterization of Model Physics Uncertainty for Simulations of Deep Moist Convection. Mon. Wea. Rev., 138, 1513–1535.
- Posselt, D. J., and G. G. Mace, 2014: MCMC-Based Assessment of the Error Characteristics of a Surface-Based Combined Radar–Passive Microwave Cloud Property Retrieval. J. Appl. Meteor. Clim., 53, 2034-2057.
- Posselt, D. J., 2016: A Bayesian Examination of Deep Convective Squall Line Sensitivity to Changes in Cloud Microphysical Parameters. J. Atmos. Sci., 73, 637–665.
- Posselt, D. J., J. Kessler, and G. G. Mace, 2017: Bayesian retrievals of vertically resolved cloud particle size distribution properties. J. Appl. Meteor. Clim., 56, 745-765, doi: 10.1175/JAMC-D-16-0276.1.
- Posselt, D. J. and C. H. Bishop, 2018: Nonlinear Data Assimilation for Clouds and Precipitation using a Gamma-Inverse Gamma Ensemble Filter. Q. J. Roy. Meteor. Soc., 144, 2331-2349. https://doi.org/10.1002/qj.3374
- Posselt, D. J., F. He, J. Bukowski, and J.S. Reid, 2019: On the Relative Sensitivity of a Tropical Deep Convective Storm to Changes in Environment and Cloud Microphysical Parameters. J. Atmos. Sci., 76, 1163–1185, https://doi.org/10.1175/JAS-D-18-0181.1
- Posselt, D. J., B. D. Wilson, R. L. Storer, D. Tropf, G. A. Duffy, M. Lebsock, V. Lall, N. Niamsuwan, and S. Tanelli, 2021: A Science-Focused, Scalable, Flexible Observing System Simulation Experiment (OSSE) Toolkit, 2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Accepted.
- Storer, R. L., and D. J. Posselt, 2019: Environmental Impacts on the Flux of Mass Through Deep Convection. Q. J. Roy. Meteor. Soc., 145, 3832-3845. https://doi.org/10.1002/qj.3669.
- Zeng, X., R. Atlas, R. J. Birk, F. H. Carr, M. J. Carrier, L. Cucurull, W. H. Hooke, E. Kalnay, R. Murtugudde, D. J. Posselt, J. L. Russell, D. P. Tyndall, R. A. Weller, and F. Zhang, 2020: Use of Observing System Simulation Experiments in the U.S. Bull. Amer. Meteor. Soc., 101, E1427–E1438. https://doi.org/10.1175/BAMS-D-19-0155.1